Deflationism in Logic, Formal Theories of Truth and Philosophy

by

María Camila Collazos Chaves

Supervisor: Graham E. Leigh

Examiner: Fredrik Engström

A thesis submitted in fulfillment of the requirements for the Master in Logic, 30 credits

Date of examination: May 23th, 2023

Department of Philosophy, Linguistics, and Theory of Science

Gothenburg University

Abstract

The concept of truth has been studied and used along the areas of philosophy, logic, and mathematics. Although is the same word ("truth") it has been used in very different ways in each area. Logicians have studied truth in the context of formal theories, while philosophers have examined it in natural languages and the world in general. The Deflationist Theory of Truth, however, seems to be a middle ground between these approaches. According to this theory, truth has a thin nature, meaning that it has little metaphysical content and does not add much to a sentence when used. This work aims to explore how these two approaches can be connected and how deflationism can be understood in them. It specifically addresses the question of conservativity in the formal theory known as the Compositional Theory of Truth and how it relates to deflationism. This project includes a review and critique of the literature on the subject, including its history, logic, results, and the arguments made by its main proponents (such as Tarski, Ramsey, Halbach, Leigh, Field, Horsten, among others). Additionally, this work suggests possible solutions to the questions posed above.

The purpose of this thesis was to establish a connection between formal theories of truth, specifically axiomatic theories of truth and the theory of $\mathcal{CT}$, and the philosophical approach to truth, particularly deflationism in formal and natural language. My main contributions and results are presented in Chapter 5.2 titled "Is truth deflationist?" In this chapter, I propose two distinct labels for different forms of deflationism, address the question of whether $\mathcal{CT}$ is an appropriate theory for deflationism, with a particular focus on the conservativity feature, and suggest how philosophical and formal approaches can collaborate with each other. This entails considering how the truth predicate behaves in various formal theories when contemplating the nature of truth, as well as taking into account other philosophical positions such as pragmatism or coherence.

Acknowledgement

I want to thank and recognize Graham, supervisor of this work, for his support, advice, patience, time and for believing in me. This work could literally not be done without his role.

I also want to recognize the presence of the entire logic group in the department (Fredrik, Bahareh, Martin, Rasmus, Peter Olsson and all of the master and PhD students) during this process. Thank you for the coffee breaks, for making me feel welcome, and for making my academic journey also human. I especially want to thank Aleks and Yoann for making me happy with their silly jokes and for being there whenever I needed support. I am also grateful to all the people I met in Gothenburg (Don Carlos, Doña Maria, Juan Carlos, Andrea, etc.) for helping me adapt to this new environment. Your support made this possible.

Quiero darle las gracias a toda mi gente en Colombia. A mis padres por su infinito amor y apoyo, y a todos quienes me hicieron sentir acompañada y con quienes descubrí que ni siquiera un océano de por medio podía en contra de nuestra unión.

And last, but not least at all, thanks to Ludvig for the company, support, love, and happiness.

Declaration

I declare that the thesis here submitted is original except for the source materials explicitly acknowledged and that this thesis as a whole, or any part of this thesis has not been previously submitted for the same degree or for a different degree.

I also acknowledge that I have read and understood the Rules on Handling Student Academic Dishonesty and the Regulations of the Student Discipline of the University of Gothenburg.

Table of Contents

La verdad es que no hay una verdad,
leí en una pared de la ciudad.

_____

(de Nos, 2012)

To discover truths is the task of all
sciences; it falls to logic to discover
the laws of truth.

_____

(Frege, 1956)

La verdad es una mirada sincera.

_____

(Mosquera Restrepo, 2022)

One of the aims of this book is to
function as an inroad to the
logico-philosophical literature on
truth.

_____

(Horsten, 2011)

1

Chapter 1

Introduction

The study of truth is concerned with understanding what it means for something to be true. This concept has been explored in various areas such as logic, philosophy, and mathematics. Philosophy has a long history of examining the nature of truth, with Aristotle being an early contributor to the discussion. In contrast, the study of truth in mathematics began with the work of Gödel, who raised questions about truth in arithmetic and incompleteness in formal systems. This led to the development of Formal Theories of Truth, which aim to provide a rigorous and systematic understanding of truth within a formal framework. Overall, the concept of truth is complex and multifaceted, and its study continues to be a topic of active research and debate in various fields.

When considering the relationship between the concept of truth in the philosophical tradition and in mathematical logic, I do not see a clear connection. The two positions, or ways of approaching the problem of the concept of truth, seem to be completely disconnected. Is it even possible or interesting to consider a closer relationship between the two?

The motivation for this study arises from the intuition that, in general, despite sharing many common problems and concepts, there is often a disconnect between philosophy and mathematics. It is important to bridge this gap in order for both disciplines to learn from one another. Additionally, this study seeks to deepen our understanding of the concept of truth, which is a fundamental concept in the philosophy of language, logic, and mathematics. As such, the goal of this work is to create a connection and contribute to building a bridge between logic and philosophy, specifically with regards to the two positions mentioned previously.

The purpose of this thesis is to establish a connection between formal theories of truth, specifically axiomatic theories of truth and the theory of $\mathcal{CT}$, and the philosoph-

ical approach to truth, particularly deflationism in formal and natural language. My main contributions and results are presented in Chapter 5.2 titled "Is truth deflationist?" In this chapter, I propose two distinct labels for different forms of deflationism, address the question of whether $\mathcal{CT}$ is an appropriate theory for deflationism, with a particular focus on the conservativity feature, and suggest how philosophical and formal approaches can collaborate with each other. This entails considering how the truth predicate behaves in various formal theories when contemplating the nature of truth, as well as taking into account other philosophical positions such as pragmatism or coherence. Furthermore, the differentiation between two types of deflationism and the assessment of truth being deflationist in one of them will prove valuable in addressing the connection between the nature of truth in logic and truth in the world. In other words, this distinction is based on the differentiation between truth in logic and truth in the world. Consequently, from my perspective, the deflationism concerned with the success of truth in logic can be seen as a step forward in exploring the broader deflationist position that focuses on truth in the world. However, it is important to note that this text does not yet delve into the transition from truth in logic to encompassing the entirety of truth in the world. Lastly, another suggestion would be to consider the other positions in the classical debate when examining how truth behaves in formal truth theories.

The truth predicate can be used in different contexts, including natural language and mathematical language. The former refers to the languages used by humans in any context, such as English, Spanish, Swedish, etc., while the latter is a more restricted and precise language used in logic and mathematics. It is interesting to note that some logicians, such as Tarski who is a precursor in this field, draw a clear distinction between logic and philosophy, and this will have implications for the study of truth:

> Although my investigations concern concepts dealt with in classical philosophy, they happen to be comparatively little known in philosophical circles, perhaps because of their strictly technical character. (Tarski, 1944).

As we can see, according to Tarski, the formal results about the concept of truth are not known in the philosophical circles due to their formal nature. However, he believes that the concept of truth should be treated as clear and concrete as possible in

the mathematical field. As we will see later, formal theories are still not widely used in philosophical literature, which supports my intuition about the existing gap between logic and philosophy, but also motivates us to explore ways to bridge this gap or at least to investigate the reasons behind its existence. This last sentence is worth explaining in more detail. In Chapter 5.1, it can be observed that there are logicians and philosophers who have been studying the philosophy of formal truth theories, specifically focusing on the nature of truth within that restricted area. Hence, the gap is not a completely open abyss. However, my assertion in that sentence pertains to the more traditional literature, which will be examined in the second chapter. Those philosophers are concerned with the nature of truth in the world and in natural language, while omitting the formal results. Nevertheless, there is one exception that should be noted, namely "Tarski's formal theory of truth". This position is discussed in Section 2.1 and is considered a philosophical stance on the nature of truth by some philosophers like Walker (2017). However, despite considering this position, they do not take into account the formal results that have emerged since Tarski's contributions (the ones that will be presented in Chapter 3). Therefore, I maintain the statement in question.

It is important to bear in mind that throughout this text, I will discuss both the nature of truth and the truth predicate. Although these two concepts are very close and similar, they are not the same. Please note that the truth predicate is a component of the overall concept of truth. It is also part of the response that philosophers concerned with truth in the world seek. On the other hand, in formal theories of truth, the focus is on the truth predicate. This can be confusing since both formal theories of truth and the literature that examines the philosophy of these theories may also employ the term "the nature of truth". However, the implications of this usage can vary depending on who is using it and how. I reiterate, though, that I use it in the sense that the nature of truth is broader and pertains to the question "what is truth?" while recognizing that the truth predicate is an important component of the nature of truth.

Another important distinction to keep in mind before delving into the main topic is the distinction between object-language and meta-language. In short, the object-language is the language we restrict and study, the truth-predicate is then used onto that language, "is the subject- matter of the whole discussion; the definition of truth which we are seeking applies to the sentences of this language" (Tarski, 1944). Mean-

while, the meta-language is the language we use to talk about the object-language, in this case, the language in which we construct the truth-predicate. Tarski emphasizes that the terms "object-language" and "meta-language" are relative and depend on the context. What is considered the meta-language in one context may be considered the object-language in another context, which then requires the introduction of a new meta-language for the former. This creates a hierarchy of languages for the truth-predicate.

It is important to note that this distinction is significant in both the formal and philosophical approaches. In the former, two distinct languages are defined, whereas in the latter, it is a feature that one must consider when contemplating the concept of truth in natural language.

It is not possible to cover all the philosophical positions on the concept of truth and all formal theories in logic in this discussion. Therefore, I will focus on the deflationist position in philosophy, which is widely accepted, and how it can be related to some formal theories of truth. Specifically, I will examine how deflationism relates to the feature of conservativity and how some logicians have attempted, and possibly failed, to establish a connection between the two.

Deflationism about truth is generally understood to be the view that truth is a "thin" notion in the sense that it lacks a deep meaning. Philosophers such as Frege have compared sentences like "the snow is white" and "it is true that the snow is white" and concluded that they mean the same thing. Deflationists would thus argue that adding the truth predicate to any sentence does not add any meaning, and that truth itself is not substantial.

Two questions that arose from this investigation were whether the compositional theory ($\mathcal{CT}$) sufficiently characterizes[1] the deflationist position (i) and whether the logical results about truth can be extrapolated to the more general and philosophical questions about the concept (ii). In this regard, I will argue that $\mathcal{CT}$, in some sense, characterizes the deflationist position, but with certain nuances that I will clarify later.

---

[1]In the literature on the philosophy of formal truth theories, there is often mention of a truth theory ($\mathcal{CT}$, in this case) that "characterizes" the deflationist position. This statement can be interpreted in various ways. In this text, my interpretation is that the appropriate truth theory embodies or represents what the deflationists understand as the concept of truth. While I am inclined to answer "yes" to that question, I will also provide a more detailed assessment as part of the contribution of this work in Chapter 5.2.

Additionally, I will argue that while the results of formal theories can be useful for future philosophical investigations about truth, it is not appropriate to assume that the logical results imply that truth is deflationist in a more general philosophical perspective. Although these are two different questions, they are related in the sense that there is an unmentioned assumption or gap in the literature regarding the two ways in which the deflationist position is understood.

1.1 The distinction

Although the justification for this distinction was hopefully implicit in the preceding lines, it is still important to explain it explicitly here. Firstly, both the logical approach to truth (Truth in Logic) and the philosophical approach (Truth in the World) aim to understand the concept of truth. However, it is clear that these two approaches have distinct motivations and methodologies for achieving their objectives.

As we have seen, philosophers tend to be much more ambitious when trying to understand the concept of truth, as they are concerned with natural languages and attempt to consider truth in a wider range of contexts. They may also reflect on their own intuitions about the nature of truth. In contrast, the formal theory approach is more specific as it restricts the language to a formal one and clearly distinguishes between the object-language (usually $\mathcal{PA}$) and the metalanguage, and focuses in the truth predicate.

Due to these differences, it is useful for this work to maintain this distinction because it helps to recognize the existing gap between how these two approaches reflect on the concept of truth and, in particular, what is the difference within the deflationist positions. However, this distinction can become blurry, particularly when considering Tarski's biconditionals as will be shown, and when some deflationist theories are exposed, as we will see in the following sections. Nevertheless, the difference in methodologies and how these two approaches use Tarski's contributions will be evident. Moreover, this distinction helps to justify the questions that I will focus on in the next section, such as how the two approaches co-exist and nurture each other, what the gap is between them, and how a bridge could be established among them. I will also explore whether a theory like $\mathcal{CT}$ characterizes deflationism.

1.2 The liar paradox

One important issue related to the concept of truth, both from a philosophical and technical perspective, is the liar paradox. An informal way to approach this matter is to consider the statement "I am lying". This statement has been considered a paradox because if it is true that the person uttering it is lying about it, then it is also stating the truth. However, if it is stating the truth, then the person is not lying.

This paradox formalizes well-known paradoxs referred to in natural language as the liar paradox, where someone says "what I am saying right now is not true", as mentioned above. If the person is lying, the sentence becomes true, leading to a loop or paradox.

The main point of this paradox lies in the issue of self-reference, both from a logical and philosophical perspective. However, it is important to note that the difference between these two approaches lies in the context in which the paradox is formulated. In logic, we specify the particular theory (Theory of Expressions, Peano Arithmetic, etc.), while in philosophy, we seek to understand how this operates in any natural language. Thus, we can assume that the statement, albeit with many variations, in the philosophical literature becomes:

(ii) This sentence is false.

From a philosophical perspective, it is not even clear if the sentence (ii) is actually referring to sentence (ii). This ambiguity has been noted by some philosophers. Another approach to the paradox, as will be mentioned in the "Truth in the world" section, involves questioning the reference in sentences that contain the truth predicate and whether such reference necessarily leads to a paradox. This can be found in (Frápolli and Nicolás Marín, 1997), (Picollo, 2020a) and (Picollo, 2020b).

The liar paradox plays a significant role in the understanding of truth, as it acts as a limit or boundary that logicians have had to address in order to define consistent theories. Furthermore, certain philosophers (Jago, 2018) have employed the liar paradox as an objection against deflationist views of truth. Whether or not their arguments are convincing is a separate matter. For further in-depth formal details, please refer to

(Halbach and Leigh, 2023) [2].

1.3   The strategy

The strategy to tackle this matter, in other words, the plan of this work is as follows: In the former chapters, I will provide the necessary background for both the formal theories of truth (which I refer to as "Truth in Logic") and the philosophical discussion of truth and deflationism (which I refer to as "Truth in the World"). For the formal theories, I will introduce the basic definitions required to understand this area of logic. This entails comprehending the distinction between the semantic and axiomatic approaches, with the Compositional Theory being defined within the axiomatic approach. Some important concepts, as conservativity, will also be introduced. In the philosophical discussion, I will explain the "Classical" and "Modern" debates, both of which are essential to understand the current state of the art in which the concept of truth is studied in these two different contexts.

In the second part of this work, I will delve into the concept of deflationism itself and introduce two labels that differentiate between two types of deflationism that I believe should be kept in mind. I will present some formal results from the truth theories, as well as how some logicians and philosophers have approached the problem and their positions. Finally, I will provide my own analysis of these positions and state my stance on the matter. As mentioned briefly earlier, I believe it is crucial to consider both positions of deflationism ($Deflationism_W$ and $Deflationism_L$), acknowledging that the arguments in one position may seem convincing and contribute to the other position, particularly in the context of the controversy surrounding conservativity. Simultaneously, we should investigate the concepts of "explanation" and "justification", as suggested by some logicians. Additionally, we should continue exploring how truth behaves in formal theories, considering deflationism as well as other classical positions. It is important to note that this aspect has not been thoroughly addressed in the literature I have examined. Furthermore, the shift towards a deflationist position in natural language remains an open question.

It is important to clarify that the selection of those philosophers and logicians in

---

[2]I would like to express my gratitude to Leigh for sharing the draft version of the book, which will be published soon.

the chapter "Closing the Gap" is based on both their significant contributions to formal theories and their engagement with the philosophy of truth in logic. Additionally, the decision to focus on $\mathcal{CT}$ is influenced by the attention it has received in the literature when exploring deflationism.

Chapter 2

Truth in the world

Philosophy is concerned with what can be said about the world or what exists, which includes the concept of truth itself. This makes it one of the broadest topics in philosophy. Moreover, truth can be studied from different areas of philosophy, such as metaphysics. In metaphysical discussions, the central question concerns the nature of truth or the essence that things we attribute truth share.

Other questions regarding the concept of truth and its implications include the problem of (anti)realism and the relation of truth to the theory of meaning. Realism asserts that the world exists objectively, independent of our thinking or description of it, and that our thoughts are about the world (Glanzberg, 2021). Realists claim that some statements are "objectively" true or false and that language accurately captures the world. However, philosophers such as Dummet reject this position and propose an anti-realist view of the world, in which they also reject a classical approach to truth and propose a logic in which a sentence can lack a value of truth and be indeterminate. Part of the intuition behind this position is that truth is determined by our abilities or what they call our "epistemic situation". On the other hand, the theory of meaning offers a view that relates the truth predicate to the meaning in natural language, either to reject or support it (Dummett, 1976).

Now, the question about the nature of the truth predicate, or truth in general, has a long history that can be traced back to the Ancient Greeks. Indeed, different philosophers and logicians (Horsten, 2011) quote the following from Aristotle:

> To say of what is that it is not, or of what is not that it is, is False, while to say what is that it is, or what is not that it is not, is true.

> Let us next say what this principle is: that it is impossible for the same thing both to belong and not to belong at the same time to the same thing and in the same respect (and let us assume we have drawn all the further

distinctions that might be drawn to meet logical complaints). (Aristotle, 1924).

Nonetheless, I will now discuss two significant portions of the debate, which represent the latest developments in a chronological approach. The first portion is what I refer to as the "classical debate", which encompasses the main positions from the twentieth century. Leading philosophers such as Russell, Wittgenstein, Ramsey, Peirce, Frege, and Moore have contributed to these theories. According to Glanzberg and Walker (Glanzberg, 2021; Walker, 2017), these positions can be summarized into five categories: the correspondence theory of truth, the coherence theory of truth, the pragmatic theory of truth, Tarski's formal truth theory, and the redundant theory of truth.

In the modern debate, two main issues are at the forefront. The first issue involves the dichotomy of whether the truth predicate is substantial or not. The second issue pertains to the question of whether the truth predicate requires a truth-maker and what the nature of a truth-maker is.

Understanding the background is crucial for comprehending the context of deflationism, its relation to other positions, and the implications of the classical debate on the modern and current discourse. It is worth noting that the redundant theory of truth in the classical debate aligns with deflationism, while the insubstantial theory of truth in the modern debate corresponds to deflationism.

Dividing the debate into these two parts is justified by the evolving labels and the increased precision in addressing specific concerns. This division reflects how the debate has transformed over time, allowing for a more focused exploration of the issues at hand. Moreover, it is important to keep in mind that these two labels are not derived from other philosophers but are suggested within this text. The shift that I have observed in the literature is based on the fact that in the classical debate, different labels and interests were associated with each position, creating a sense of divergence between them, despite all of them contemplating the nature of truth in natural language. In contrast, philosophers engaged in the modern debate appear to have two primary concerns regarding the nature of truth in natural language: the truth-bearers and the characteristic of (in)substantiality of truth, with the latter pertaining to the deflationist position.

2.1 The classical debate

The ordering of the five positions may vary depending on different perspectives in the literature, with one possible arrangement being provided in (Glanzberg, 2021). However, another approach distinguishes the classical debate into these five categories, labeling the last theory as the "Semantic theory of truth" (Walker, 2017). I will explain why I did not adhere to that specific label.

Each of these five theories attempts to address the question of what constitutes the nature of truth in its own way. It is important to note that the redundancy theory of truth is closely associated with, or even gives rise to, the deflationist theory of truth.

In the classical debate, the correspondence theory of truth holds that the truth predicate corresponds to the way things actually are in the world. This theory, advocated by Russell and Moore, suggests that when we use the truth predicate, we are asserting a proposition that corresponds to a fact. For example, the proposition "it is true that it is not that cold in Gothenburg anymore" corresponds to the fact that the temperature in Gothenburg these days is above 8 degrees. According to this theory, the bearers of truth are beliefs rather than propositions. The slogan associated with this view is: "A belief is true if and only if it corresponds to a fact" (Glanzberg, 2021).

There are different variations within the correspondence theory. One such variation is the identity theory, which posits that a true proposition is identical to a fact. For instance, the proposition "it is true that Ana is talking" corresponds to the existence of the individual Ana and the property of talking. As there is a fact in the world that matches this proposition, it is considered true. The distinction here lies in separating individuals from their properties.

Wittgenstein's truth theory provides another perspective within the correspondence theory. According to Wittgenstein, true propositions have a one-to-one correspondence with the world. He argues that there is an isomorphism between language and the world, where the logical structure of language corresponds to the logical structure of the world. This correspondence is established because of the shared logical structure between the two (Wittgenstein, 2009).

It is worth noting that the correspondence theory of truth offers various ap-

proaches and interpretations, but they all share the fundamental idea that truth involves a correspondence between propositions and the way things are in the world.

> To the configuration of the simple signs in the propositional sign corresponds the configuration of the objects in the state of affairs. (3.21). (Wittgenstein, 2009).

The coherence theory of truth holds that truth is determined by the coherence of a belief within a larger system of beliefs. According to this view, a belief is considered true if and only if it is consistent and logically entailed by the other beliefs in the system. The focus is on the internal consistency and logical coherence of the beliefs, rather than their correspondence with the external world.

The coherence theory is often associated with idealism, particularly British idealism. It suggests that truth is not primarily dependent on how the world justifies beliefs, but rather on how the beliefs are interconnected and mutually support each other within a coherent system. Some coherence theorists argue that the coherence should be understood in terms of what one would believe under idealized or optimal circumstances, rather than what people actually believe.

Immanuel Kant can be seen as a coherence theorist, although his version of coherence theory also included the consideration of the "material conditions of experience", which encompassed sensory input and empirical data (Walker, 2017). Kant emphasized the importance of the systematic coherence of our concepts and judgments in relation to our experience of the world.

In summary, the coherence theory of truth emphasizes the internal consistency and logical coherence of beliefs within a system, and it is often associated with idealism and the notion that truth is determined by the relationships among beliefs rather than their correspondence with the external world.

The pragmatic theory of truth asserts that the nature of truth is determined by its practical usefulness or effectiveness. According to this view, a belief is considered true if it proves to be beneficial or successful when put into action. The main proponents of this theory are Charles Peirce, John Dewey, and William James.

The pragmatic theory suggests that the truth of a belief is evaluated based on

whether it "works" or leads to the expected outcomes when acted upon. The criterion for truth is the coherence between the belief and future experiences. If the belief accurately predicts and aligns with our future experiences, then it is considered true. This coherence with future experience is what determines the truth value of the belief.

The pragmatic theory of truth emphasizes the practical consequences and utility of beliefs. It views truth as a product of human inquiry and experimentation, rather than an independent and objective reality that exists apart from human propositions and beliefs. The focus is on the instrumental value of truth, as it guides our actions and decisions in pursuit of successful outcomes.

In summary, the pragmatic theory of truth posits that truth is determined by the practical usefulness and effectiveness of beliefs. It emphasizes the coherence between beliefs and future experiences, and it views truth as a result of human inquiry and experimentation. The focus is on the practical consequences and benefits of holding true beliefs.

The origins of the redundancy theory can be traced back to Frank Ramsey, who proposed the equivalence thesis (Ramsey and Moore, 1927). According to Ramsey, the statement "p is true" is equivalent in meaning to simply stating "p" itself. This idea aligns with Alfred Tarski's biconditional account of truth, which states that a statement is true if and only if it corresponds to the facts.

In the redundancy theory, propositions are considered as the bearers of truth, and there is no need to posit a separate property of truth. This approach focuses on the truth predicate itself and its relation to propositions.

It is important to note that there is no inherent conflict between the redundancy theory (or deflationism) and the other theories discussed earlier. The redundancy theory specifically addresses the nature of the truth predicate, while the other theories examine the broader concept of truth. However, this raises the question of why there is a distinction when we use the truth predicate in relation to the broader concept of truth.

In summary, the redundancy theory posits that the concept of truth is redundant, and theories influenced by this view aim to deflate the significance of truth. The focus is on the truth predicate and its relation to propositions, with no separate property of

truth. Deflationist theories seek to preserve the essence of the redundancy theory while making necessary adjustments.

Finally, Tarski's formal theory of truth holds a significant position in this debate. Like Ramsey, Tarski proposes a biconditional account of truth and presents a crucial definition. He introduces a convention, known as convention T, which establishes an adequacy condition for formal theories of truth. Tarski diverges from the view that propositions or beliefs are the bearers of truth and instead considers sentences as the bearers. While some philosophers, such as Walker, classify Tarski's theory as a semantic theory of truth, I have refrained from using that label since it is employed in contemporary logical literature to denote a more precise and distinct theory, as will be explained later.

According to Walker, Tarski formulates a set of axioms for the language and distinguishes between the object language and the metalanguage to prevent paradoxes. Thus, there are restrictions on the use of truth predicates applied to themselves. This context can be extended to natural language, where the axioms of the object language correspond to certain entities or sets of entities in the world. For example, "Tarski" would correspond to Tarski. This relationship, in a rudimentary manner, is understood as satisfaction. However, Walker argues that this theory does not provide a comprehensive account of truth itself and, therefore, fails to fulfill our quest for understanding. He states, "this gives us no general account of what truth consists in, and so it does not give us what we were looking for".

In fact, Tarski's intention was to study the truth predicate solely within formal languages, as he believed it was futile to investigate it in broader contexts as other philosophers attempted. The reason for this lies in the fact that truth concerning the object language is defined in the metalanguage, necessitating a universal definition of truth for all languages, which is unattainable.

2.2 The modern debate

In contrast to the previous subsection, the current debate centers around two key concepts: truth-bearers and (in)substantialism. In the contemporary discourse on the metaphysical nature of truth, there are three primary positions, as outlined by Jago (Jago,

2018). One of these positions, known as substantialism, posits that truth is a property inherent to entities. These entities are propositions, which represent the content conveyed by declarative sentences. According to substantialism, when making an assertion such as:

(i) Tarski is the greatest logician.

I am asserting the proposition that "Tarski is the greatest logician," and this proposition does not belong to any particular language such as English, Spanish, or Swedish. Instead, it represents the content that all of those languages express when positively uttered. According to substantialists, a proposition ϕ is considered true only if there exists something that serves as its truth-maker. This "something" is commonly referred to as the "truth-maker" in their framework. Different labels may be used for truth-makers; for instance, Jago refers to them as "states-of-affairs", while Mulligan uses the term "moments" as an example (Mulligan et al., 1984). Jago justifies the existence of truth-makers by pointing to the existence of properties and relations in the world. When a proposition is deemed true, it indicates that there is something within the world that accounts for its truth.

On the other hand, the insubstantialist (or deflationist) position argues that the truth predicate is insubstantial. It denies the metaphysical commitment to a truth-maker when using the truth predicate. The cornerstone of this position is the Tarski T-schema:

(ES) <p> is true if and only if p.

This is how it is stated in (Armour-Garb et al., 2023), where "ES" stands for what Armour-Garb and Woodbridge refer to as the "equivalence schema". The brackets serve as an appropriate designation, akin to quotation marks. For a deflationist, this schema is fundamental in understanding the truth predicate. Furthermore, there are three distinct versions of deflationism in philosophy: sententialism, propositionalism, and prosententialism. The first version asserts that the instances of the (ES)-schema pertain to sentences, the second version asserts they pertain to propositions, and the third version introduces a new concept called prosentences, based on Ramsey's ideas. Although I will not delve into each of these versions here, it is still important to acknowledge their existence and significance in the ongoing debate. However, while

further exploring this aspect of the investigation would have been intriguing, the decision not to focus on it stems from the breadth of the topic concerning the nature of truth, encompassing both philosophical and formal approaches. Thus, selectivity was necessary, which meant not delving into all the issues found in the literature. Starting from Chapter 4, I have adopted a deflationist position that has been employed in the literature examining the philosophy of formal theories, albeit specifically within the context of formal languages. Nonetheless, it is hopeful that this aspect will be explored in the future.

Although we can conclude from the previous section that this version of deflationism originated from Ramsey's redundancy theory, Dummett argues that it actually began with Frege. He states:

> It is ... worthy of notice that the sentence 'I smell the scent of violets' has just the same content as the sentence 'It is true that I smell the scent of violets'. So, it seems, then, that nothing is added to the thought by my ascribing to it the property of truth. (Frege 1918, 6).

But, in any case, the influence of Ramsey is undeniable. One important contribution is how Ramsey recognizes that:

> truth predicate fills a grammatical need, which keep us from eliminating it altogether, but he held that even in these cases it contributes no content to anything said using it. (Armour-Garb et al., 2023).

This leads us to the feature of generalization and the generalization problem, which is still being discussed by the main proponents in the present, as we will explore in subsequent chapters. Essentially, this problem revolves around the notion that although truth is thin, it remains useful for expressing generalizations.

Finally, we should consider other versions of deflationism, namely minimalism (mainly advocated by Horwich) and disquotationalism. Minimalism aligns with substantialists in the sense that the truth predicate describes entities[1], but it diverges by

---

[1]To describe entities means to provide information or attributes that characterize or define specific objects, individuals, or concepts. Describing entities involves conveying relevant details about their characteristics, properties, qualities, or attributes in order to provide a clear understanding.

The process of describing entities typically involves identifying and expressing key features or aspects that distinguish them from others. This could include providing information about their appearance, functional attributes, behavior, relationships, or any other relevant characteristics.

asserting that the predicate itself does not possess a substantive nature. Instead, its function lies in generalizing propositions (Armour-Garb et al., 2023). On the other hand, disquotationalism rejects the notion that the truth predicate describes entities. This position, primarily championed by Field, is based on the Tarski schema and also acknowledges the generalization aspect of the truth predicate.

Since deflationism is the central focus of this investigation, I will delve deeper into it later. For now, I would like to clarify that the preceding discussion provides a more philosophical rather than formal summary of what deflationism entails. However, one challenge we will encounter when studying the philosophical approach is the lack of a universally accepted or clearly defined definition of deflationism. Furthermore, we will explore how certain logicians, such as Halbach, approach this position in subsequent chapters.

Lastly, the third position in the modern debate is pluralism, which essentially combines both the substantialist and insubstantialist positions by asserting that the nature of the truth predicate can vary and lacks homogeneity. It is also worth noting that deflationism is not the sole approach aimed at challenging the substantialist position. There are other positions that simply deny the existence of facts (or states of affairs or truth-bearers) without specifically addressing the issue of truth. For further exploration of these topics, please refer to (Betti, 2015).

In summary, describing entities involves articulating relevant information about their defining characteristics or attributes, enabling others to form a clearer understanding or mental representation of those entities.

Chapter 3

Truth in Logic

In this section, I will introduce what I refer to as "Truth in Logic". In contrast to the previous section, there is a restriction: we focus solely on formal languages and set aside considerations of natural languages. Consequently, we no longer contemplate the concept of truth in the general world, but rather we are specifically concerned with the concept (or predicate) of truth in logic. The theories of the following sections, "Semantic truth" and "Axiomatic truth", are commonly referred to as formal theories of truth in the literature. It is crucial to understand these theories because, according to several philosophers and logicians, they have a strong connection to the deflationist position. This suggests that these theories could provide support for the deflationist understanding of truth. However, as we will see, this position is subject to debate and requires careful examination. Furthermore, another justification for delving into this topic is that when one seeks to comprehend the concept of truth, it can be advantageous to work with a restricted and formal language rather than starting from scratch with natural language.

In principle, there are two approaches to tackling the problem regarding the nature of truth in formal theories: the model-theoretic (or semantic) approach and the axiomatic (or syntactic) approach. In the model-theoretic approach, the notion of truth is a defined concept. Conversely, in the axiomatic approach, we have a theory that encompasses a set of axioms aiming to elucidate how the concept of truth operates. In this approach, the notion of truth is incorporated into an extended language with a one-place predicate. Throughout this section, the basic definitions of are taken from (Halbach, 2014), (Cieśliński, 2017), and (Horsten, 2011). Each logician provides its own unique version, but they ultimately refer to the same theories. Here, I present them in a manner that I believe to be the clearest and most practical.

The following is a foundational and very important criterion suggested by Tarski

that influenced both semantic and syntactic truth[1].

Convention T: *A formally correct definition of the symbol "T" is an adequate definition of truth if it has the following consequences:*

(α). *All sentences obtained from the expression 'x $\in$ T if and only if p'. And where 'x' is a structural descriptive name of any sentence of the language and 'p' the expression which forms the translation of this sentence into the extended language.*

(β). *The sentence 'for any x, if x $\in$ T then x $\in$ Sent' (T $\subseteq$ Sent).*

The first condition uses a "structural descriptive name" to refer to what will be the quotation of the sentences in the following. For example, in $\mathcal{PA}$, it would be the numeral code. The method involves listing each letter of the English alphabet and then creating a structural descriptive name for a word like "Tarski" by listing the corresponding numbers in the alphabet. Now, the second condition states that only sentences can be true.

It is important to emphasize the significance of Convention T in the overall discussion, not only in the study of truth in logic but also in understanding truth in the world, as briefly mentioned in the last section. Firstly, this definition is credited to Tarski and it revolutionized how philosophers and logicians approached the concept of truth and the truth predicate. Tarski formulated the convention with the aim of studying truth in a more precise manner compared to the prevailing approaches of that time. Additionally, his intention was to propose a semantic theory of truth, although it is intriguing to see how this convention ended up influencing the development of axiomatic theories of truth. It is also attributed to Tarski for defining truth in a metalanguage for an object-language and for making a clear distinction between different languages. However, it is important to note that this convention does not actually provide a definition of the concept of truth; rather, it offers an adequacy criterion for assessing any proposed definition of truth.

---

[1]As mentioned in the "Truth in the World" section, Tarski and his convention T are part of the ongoing debate. This doesn't imply that he is closing the gap, but rather that philosophers, at least up to that point in the narrative, were aware of what logicians were working on with regard to truth and even considered it in their philosophical discussions.

3.1 Preliminaries

In the following I present some preliminary definitions that are necessary to understand the following two sections and the section 4.2. Recall that a language $\mathcal{L}$ contains a set of function symbols, a set of relation symbols and a set of constant symbols. Also, that a structure $\mathcal{M}$ in the language $\mathcal{L}$ contains a nonempty domain, a function, a set of relations and constants that are the "interpretations" of the correspondent symbols in the language. The definition of a theory, as taken from (Halbach, 2014), is that it is a set of formulas closed under first-order logical consequence. Therefore, a theory can be generated by many different formal systems.

Definition 1 L_{PA} *denotes the language of first-order arithmetic and it contains quantifiers, connectives, brackets and variables as usual, and the extra-logical symbols $\{`+', `\cdot', `0', `S'\}$ which correspond to adition, multiplication, zero and successor function.*

Definition 2 $\mathcal{PA}$ *is the theory with the following axioms:*

1. *$\forall x \forall y (x' = y' \rightarrow x = y)$.*

2. *$\forall x 0 \neq x'$.*

3. *$\forall x (x = 0 \vee \exists y x = y')$.*

4. *$\forall x (x + 0 = x)$.*

5. *$\forall x \forall y (x + y') = (x + y)'$.*

6. *$\forall x (x \cdot 0) = 0$.*

7. *$\forall x \forall y (xy') = ((xy) + x)$.*

8. *$\forall x \forall y (x < y \leftrightarrow \exists z (z' + x) = y)$.*

9. *If $\phi(x)$ is a formula in the language with free variables x and $\overline{y}$, any sentence of the form: $\forall y_1 ... \forall y_n ((\phi(0) \wedge \forall x (\phi(x) \rightarrow \phi(x'))) \rightarrow \forall x \phi(x))$, then it is an instance of the induction schema and thus part of the theory.*

Definition 3 *Let $\mathcal{L}_t$ be the language $\mathcal{L}$ of arithmetic together with a new unary (truth) predicate T. And let label $\mathcal{PAT}$ for Peano Arithmetic formulated in $\mathcal{L}_t$ with all instances of the induction schema in the language $\mathcal{L}_t$.*

Definition 4 *Let Sent(x) be the set of all sentences of $\mathcal{L}$ and $Sent_t(x)$ the set of all sentences of $\mathcal{L}_t$.*

The following definition helps to understand the Compositional Theory in the section 3.3.1.

Definition 5 *If ϕ and ψ are formulas, and p_i is a propositional variable, then $\phi(\psi/p_i)$ denotes the result of replacing each occurrence of p_i by an occurrence of ψ in ϕ; similarly, the simultaneous substitution of $p_1, ..., p_n$ by formulas $\psi_1, ...\psi_n$ is denoted by $\phi(\psi_1/p_1, ..., \psi_n/p_n)$.*

The following definitions are required to prove the theorems in the section 4.2.

Definition 6 *$Prov_{Th}$ is a formula of L_{PA} that represents in PA the recursive relation 'a is a proof of γ from a fixed recursive set that axiomatizes the theory Th'. Pr_{Th} is the formula $\exists x Prov_{Th}(x, y)$.*

Definition 7 *Con_{Th} denotes the sentence $\neg Pr_{Th}(\ulcorner 0 = 1 \urcorner)$.*

Definition 8 *M is recursively saturated if and only if every recursive type of M is realised in M.*

Definition 9 *$\alpha \sim_{v_i} \beta$ is the formula $Asn(\alpha) \wedge Asn(\beta) \wedge \forall k[k \neq i \rightarrow \alpha(v_k) = \beta(v_k)]$.*

Definition 10 *If the language $\mathcal{L}_f$ is an extension of L_{PA} with a binary predicate S, $PA(S)^-$ is the extension of PA together with:*

$$\forall t, s, \alpha[Tm(t) \wedge Asn(\alpha) \rightarrow (S\ulcorner t = s \urcorner, \alpha) \leftrightarrow val(t, \alpha) = val(s, \alpha)]$$

$$\forall \psi, \alpha[Fm(\psi) \wedge Asn(\alpha) \rightarrow (S\ulcorner \neg \psi \urcorner, \alpha) \leftrightarrow \neg S(\psi, \alpha)]$$

$$\forall \psi, \phi, \alpha[Fm(\psi) \wedge Fm(\phi) \wedge Asn(\alpha) \rightarrow (S(\ulcorner \psi \wedge \phi \urcorner, \alpha) \leftrightarrow S(\psi, \alpha) \wedge S(\phi, \alpha))]$$

$$\forall \phi, v_i, \alpha[Fm(\phi) \wedge Asn(\alpha) \wedge Var(v_i) \rightarrow (S(\ulcorner \exists v_i \phi \urcorner, \alpha) \leftrightarrow \exists \beta(Asn(\beta) \wedge \alpha \sim_{v_i} \beta \wedge S(\ulcorner \phi \urcorner, \beta)))].$$

3.2 Semantic truth

In this section, I will explain the nature of the semantic approach and why it is not widely embraced among logicians. It is important to note that in this methodology,

truth is treated as a defined concept. An interpretation of a truth predicate is provided in a metalanguage (Fischer et al., 2015). This is achieved by assigning a function and establishing a satisfaction relation within a model. The specific method employed for this purpose varies depending on whether the theory being studied is classical logic or non-classical logic. If it is classical, the approach known as the "Tarski method", as referred to by (Cieśliński, 2017), is utilized. For non-classical theories, a Kripke construction is employed. To fully grasp the definition of truth within this framework, it is necessary to familiarize oneself with the following foundational definitions. (Marker, 2006):

Definition 11 *A model M of T (in symbols: $M \models T$) if $M \models \phi$ for all sentences $\phi \in T$.*

Definition 12 *If M_1 is an L_1-structure and M_0 is an L_0-structure such that M_0 ignores the interpretation of the symbols in $L_1 \setminus L_0$, we say that M_1 is an expansion of M_0 and M_0 a reduct of M_1.*

Next, we consider a family of languages that extends L_{PA} and an arbitrary model M of L_{PA} (represented as $M = (|M|, +M, \times_M, S_M, 0_M)$) with the same signature as L_{PA}. The signature specifies the set of constants, n-ary relation symbols, and the set of n-ary function symbols of the structure (Hodges et al., 1993). Within this model, we introduce an assignment, which is a function mapping variables (Var) to the domain of M. We also define a function $val^M(t, a)$, representing the value of a term t in model M under assignment a, defined recursively. Additionally, a satisfaction relation between the model, a formula ϕ, and an assignment is established recursively, with the value of terms playing a role in determining the satisfaction of atomic formulas (Cieśliński, 2017).

After that, the expression '$a \sim_\phi b$' is defined as that for every variable x free in ϕ, $a(x) = b(x)$ which means that we assume a and b coinciding on variables which are free in ϕ. Then, we have that for all a and b and ϕ, if $a \sim_\phi b$, then $M \models \phi[a]$ if and only if $M \models \phi[b]$. Thus, sentences are satisfiable by all assumptions or by none. (Cieśliński, 2017). And then, finally, the truth definition is defined as follows:

Definition 13 *For every sentence ϕ, $M \models \phi$ (which is read as 'ϕ is true in M') if and only if for every assignment a, $M \models \phi[a]$.*

This definition is the notion of truth under interpretation in a model under Tarski's method. Then, for the notion of truth simpliciter, one picks a model and define it as follows:

Definition 14 *For any structure M, and define for all sentence $\phi \in L_{PA}$, ϕ is $true_0$ if and only if $M \models \phi$.*

The '0' under 'true' indicates its place in the hierarchy, meaning that is the first notion of truth. This definition applies only to sentences of L_{PA} which do not have the truth predicate. The following is the definition of T_0.

$$(M, T) \models T_0(t)[a] \text{ if and only if } val^{(M,T)} \in T.$$

And the definition of a standard model M_1 for $L_{PA}^{T_0}$ is $M_1 = (M, (Th(M)))$. Where M is the standard model introduced before and Th(M) is the set of all formulas in the language that are true in M. Finally, the truth simpliciter is defined in this stage of the hierarchy as:

for every sentence $\phi \in L_{PA}^{T_0}$, ϕ is $true_1$ if and only if $N_1 \models \phi$.

This procedure can be iterated to extend the notion of truth simpliciter to higher levels. However, it is worth noting that the definition relies on the notion of a standard model, which poses a limitation and is one of the reasons why this approach is often disregarded.

Now, Cieśliński also mentions Kripke's construction, and it is important to briefly mention it here as well. It essentially follows the same procedure as Tarski's approach, but with the difference that it is not in the classical realm. Therefore, the notion of a model in the language is not classical, as the truth predicate is partial. In Kripke's construction we work with $\mathcal{L}_{\mathcal{T}}$ with a one-place predicate T. However, the notion of a model of that language is not classical and the interpretation of 'T' is as a partial predicate. That partial interpretation is given by to sets (the extension and the anti-extension of the one-place predicate). And, then, the elements pertaining to the extension would considered as true, the ones belonging to the anti-extension would be considered as false and the elements that do not belong to any of those would be considered as indetermined. This can be observed in the following definition:

Definition 15 *A model in the language L_T is any structure $M = (K, T^+, T^-)$. This is*

an expansion of K with the language L_{PA} where T^+ is called the 'extension' and T^-
the 'anti-extension' in M.

Note that from the start, the language being considered already contains the truth predicate, which is a difference from Tarski's approach. Also, consider that it might be the case that the union of the extension and the anti-extension is a proper subset of the set of codes of sentences in K. This is the case for partial models. Therefore, the notion of truth becomes:

If $M \models_{SK} \phi$, then ϕ is true in M;

if $M \models_{SK} \neg\phi$, then ϕ is false in M

and if neither of those holds in the model, then ϕ is considered as indetermined. Where '$M \models \phi[a]$' refers to the satisfaction relation in the partial models, this belongs to what is known as Strong Kleene Logic; see (Halbach, 2014). Then, the procedure to define the notion of truth simpliciter is analogue to the one of Tarski. However, the limitation I previously mentioned regarding the requirement for the existence of models is also applicable in this case.

Even if this semantic methodology is not the most popular, there have been attempts to relate it to the axiomatic approach (Fischer et al., 2015). In these attempts, the semantic approach is not completely disregarded, and the axiomatic approach is seen as a way to axiomatize the semantic theories. They do so with three criteria.

(Criteria 1) Initially, the concept of "resemblance" is defined. Subsequently, the assessment is made that an axiomatic system is deemed satisfactory to the degree that it bears resemblance to its corresponding semantic theory.

(Criteria 2) An axiomatic theory Σ is an adequate axiomatization of a semantic theory M if and only if a theory that proves the existence of some extension satisfying Ψ_M (a formula of second-order arithmetic) is reducible to Σ.

(Criteria 3) Let Σ be an axiomatic and M a semantic L_{PAT}-theory. Σ is an adequate axiomatization of M if and only if for all $S \subseteq \omega$:

$$(*) \ (\mathbf{N}, S) \models \Sigma \leftrightarrow (\mathbf{N}, S) \in M.$$

Even though the relationship between syntactic and axiomatic theories extends beyond

the scope of this thesis, I find it still interesting to consider. I believe it could shed new light on the topic of truth and its nature from a philosophical perspective.

3.3   Axiomatic truth

This is a more popular approach, and as mentioned before, it involves extending a language with a truth predicate that aims to demonstrate how truth operates in certain logical or mathematical theories.

Generally, we work with a theory called the "base theory" from which we extend to the truth theory we will consider. The base theory we work with for this topic is typically the theory of Peano Arithmetic. However, it should be noted that we could also work with other base theories. Nevertheless, choosing another base theory would not add interesting features or consequences related to the purpose of studying the concept of truth.

Another important distinction to keep in mind in these theories is the distinction between typed and untyped theories (Halbach, 2014). Typed theories are restricted in the sense that the truth predicate is only applied to sentences that do not contain the truth predicate itself. On the other hand, untyped (or type-free) theories do not have this restriction. Halbach suggests ways in which untyped theories can avoid inconsistency without relying on the distinction between object-language and metalanguage.

All the languages I will consider have $\neg$, $\wedge$, $\vee$, $\rightarrow$ and $\leftrightarrow$ as connectives and $\exists$ and $\forall$ as quantifier symbols. The logical axioms and rules are formulated for these connectives and quantifier symbols; they also include suitable axioms or rules for identity.

$\mathcal{PAT}$ is the theory in the language L_T, whose axioms contain those of $\mathcal{PA}$, together with all the logical axioms in L_T and all the substitutions of the induction schema by formulas of L_T.

Finally, it is also important to keep in mind that, although there are numerous axiomatic truth theories, I am only presenting the compositional theory of truth here. The reason for this choice is that, given the scope of the thesis, I needed to be selective. While other theories may also be intriguing to consider, the compositional theory appears to be the most widely recognized and studied among the logicians and

philosophers discussed in chapter 5.1.

3.3.1 Compositional Theory

The compositional theory of truth (one typed axiomatic theory of truth) arises from the convention T in which every true sentence is in the set of all sentences, trivially, by (β). And condition (α) expresses in a more simple way that Sent(a), for any a, is true if and only if a. Based on Convention T, the compositional truth theory aims to provide a systematic account of truth for complex sentences in terms of the truth of their component parts. It utilizes the principle of compositionality, which states that the meaning of a complex expression is determined by the meanings of its constituent parts and their mode of combination (see (Szabó, 2022)). In the context of the truth theory, compositionality means that the truth value of a complex sentence can be determined by the truth values of its component sentences and the way they are combined (see (Halbach, 2014)).

Definition 16 ($\mathcal{CT}^-$) *The system $\mathcal{CT}^-$ is given by all the axioms of $\mathcal{PA}$ and the following axioms:*

1. $T(s = t) \leftrightarrow (val(s) = val(t))$.

2. $\forall x(Sent_{L_{PA}}(x) \rightarrow (T(\neg x) \leftrightarrow \neg T(x))$.

3. $\forall x \forall y(Sent_{L_{PA}}(x \wedge y) \rightarrow (T(x \wedge y) \leftrightarrow T(x) \wedge T(y))$.

4. $\forall x \forall y(Sent_{L_{PA}}(x \vee y) \rightarrow (T(x \vee y) \leftrightarrow T(x) \vee T(y))$.

5. $\forall v \forall x(Sent_{L_{PA}}(\forall vx) \rightarrow (T(\forall vx) \leftrightarrow \forall t T(x(t/v))$.

6. $\forall v \forall x(Sent_{L_{PA}}(\exists vx) \rightarrow (T(\exists vx) \leftrightarrow \exists t T(x(t/v))$.

The system $\mathcal{CT}^-$ is typed as all the quantifiers ranging over sentences in the axioms 2-6 range only over sentences of $\mathcal{L}$. It is compositional since the truth of a sentence depends on semantic values of the constituents of that sentence.

$\mathcal{CT}^-$ is the result of turning how Tarski tought the propositions of truth to axioms. Then, $\mathcal{CT}$ reflects the disquotation schema of Tarski but turns the idea into these six axioms. Also, as the main idea of α remains as it works for atomic sentence and is extended to the logical connectives as we saw above.

Definition 17 ($\mathcal{CT}$) *The system $\mathcal{CT}$ is obtained from $\mathcal{CT}^-$ by adding all induction axioms in the language of $\mathcal{L}_t$.*

The following four theories are presented by Cieśliński as theories that represent the disquotationalist position since they focus on characterize truth in relation to quotation.

3.4   Conservativity

One of the formal results that has been studied by logicians interested in the philosophy of the formal theories of truth is conservativity. As we will see in the definitions and philosophical discussions, conservativity is closely related to our understanding of truth being (in)substantial. This relationship will help explain part of the debate in the literature. The definitions and proof idea are taken from (Cieśliński, 2017).

Definition 18 *Let T_1 and T_2 theories in languages L_1 and L_2 with ($L_1 \subseteq L_2$). Then:*

1. *T_2 is syntactically conservative over T_1 if and only if $T_1 \subseteq T_2$ and $\forall \psi \in L_1 [T_2 \vdash \psi \rightarrow T_1 \vdash \psi]$.*

2. *T_2 is semantically conservative over T_1 if and only if every model M of T_1 can be expanded to a model of T_2. Which means that interpretations for new expressions of L_2 can be provided in M so that T_2 is true in the expansion of M).*

To discuss this, on the one hand, notice that it is known that semantically conservativeness implies syntactic conservativeness. A brief idea of the proof consists in assuming by contradiction that T_2 is not syntactically conservative over T_1. Then, there is a sentence ϕ in L_1 such that is deduced in T_2 but not in T_1. Then, $T_1 \cup \neg \phi$ is consistent and, by completeness, there is a model where $\neg \phi$ is true. Since T_2 is semantically conservative over T_1 that model can expanded to T_2 and then, ϕ is true in T_2 and $\neg \phi$ is also true in T_2, a contradiction. On the other hand, keeping in mind what I mentioned before regarding the (in)substantiality feature of truth, conservativity appears to be a point of debate. For some philosophers, it seems that a truth theory that is not conservative cannot be considered insubstantial.

Chapter 4

What is deflationism?

In the previous chapters, I introduced the topic of truth from both a philosophical and formal theory perspective. This encompasses discussions of truth about logic and truth about the world in general. Within the philosophical approach, I provided a summary of both classical and contemporary debates. Meanwhile, within the formal theories of truth, I explained the distinction between the semantic and axiomatic theories of truth. I also outlined the main axiomatic truth theories aimed at characterizing the truth predicate[1].

As I mentioned above, it seems that everyone agrees on the main point of deflationism, which is that the meaning of truth is, in reality, unsubstantial and almost useless[2]. However, it must be mentioned that according to most logicians who have thought about truth (Halbach, 2001; Cieśliński, 2017; Horsten, 2011; Tennant, 2005; Field, 1999), there is no single or precise definition of what deflationism means.

Now, regarding the history of deflationism, it can also be said that deflationism started with Ramsey (Ramsey, 1991) and, according to some philosophers like Dummett, was inspired by Frege (Armour-Garb et al., 2023). In support of this, consider his quote (already stated in 2.2 on page 17):

---

[1]The discussion about the axiomatic theories of truth characterizing the truth predicate and/or supporting deflationism is in dispute, as we will see in Chapter 5.1. In the next section, logicians and philosophers will defend this position, which we will explore further in the next chapter. However, there is no common agreement regarding the question of whether these responses are satisfactory or not. In Chapter 5.2, I will mention that I consider the arguments presented in 5.1 to offer a good answer. Nevertheless, we should keep in mind that the discussion is still open, and even individuals like Cieśliński suggest continuing the investigation with certain concepts that will be mentioned in their own section 5.1.2.

[2]It is important to note that describing truth as "almost useless" is, at the very least, controversial and problematic. This is a viewpoint commonly held by philosophers who adopt a deflationist position, as pointed out by philosophers like Shapiro. My interpretation of this statement is that truth does not function like other properties in the sense that it does not provide new information about the object to which we attribute truth. However, there is a reason why we still use truth, specifically the truth predicate, and that reason is precisely what justifies the shift from stating that truth is completely useless to stating that truth is *almost* useless. There is no definitive answer to this matter since the nature of truth is still debated and is partly what I aim to explore in this work.

It is ... worthy of notice that the sentence 'I smell the scent of violets' has just the same content as the sentence 'It is true that I smell the scent of violets'. So, it seems, then, that nothing is added to the thought by my ascribing to it the property of truth. (Frege, 1918).

As mentioned before, he was regarded by philosophers like Walker and Jago as the main proponent of the redundancy theory of truth. Indeed, he suggested a similar version of Tarski's biconditional. In this context, the truth bearers are propositions and truth is not considered a property. According to Frápolli, Ramsey's aim was "to explain the connections between truth and falsehood, on the one hand, and the notions of judgment, belief, and assertion, on the other" (Frápolli, 2005). However, she also argues that Ramsey was not actually advocating a coherence theory regarding the truth predicate in natural languages. This is because, according to her, the biconditionals are only useful in formal languages. Nevertheless, it is agreed in the literature that Ramsey acknowledges the grammatical need for the truth predicate while simultaneously not adding any meaning to the sentences.

Now, although there is no common definition of deflationism, every deflationist adheres to Tarski's equivalence schema. Therefore, a very important feature of the truth predicate is the equivalence "p" is true if and only if p, where the quotation of p in the first part of the biconditional can be specified in different ways. According to deflationists, this equivalence explains the truth predicate and the concept of truth. This means that the equivalence shows how truth behaves and, in that sense, it responds to the question of what the nature of truth is.. For philosophers concerned with truth in the world, this equivalence is applied either to sentences or to propositions. Different versions of deflationism depend on how they interpret the schema (Armour-Garb et al., 2023). The analytic approach suggests that the truth predicate does not add new information to p. The necessary approach holds that the equivalence is a necessary biconditional, and the last position suggests that the equivalence is simply a material logical biconditional.

If we adopt the Tarski biconditional as the feature that fully characterizes how truth works (again, as mentioned in the last paragraph, this means that it shows how the concept of truth behaves, clarifying its function and utility), then truth is useless and

not a meaningful property at all. However, ascribing a property to an object modifies the object, but the truth predicate does not do so. For example, when we say "it is true that the snow is white". Thus, one way to understand deflationism is that the truth predicate is not used to ascribe any property. However, this does not mean that the truth predicate does not have any function at all.

In fact, there are other features that are recognized and mostly adopted by both philosophers and logicians. The first one is disquotationalism. This position is generally attributed to Horwich, and the version of deflationism consists of the schema mentioned above not only characterizing but also completely exhausts the truth predicate[3]. In other words, the content of truth is fully captured by that schema. Another feature recognized in the contemporary literature is that even though the truth predicate does not add any information, it is still useful because it works as a generalization device. We use the truth predicate when we want to express infinite conjunctions, for example, or in sentences like "everything that Gödel said is true".

Another area where the truth predicate might be useful is in blind ascriptions. Blind ascriptions refer to attributing a property to an object without evidence. For example, the proposition "the next thing that Ana says will be true". (Armour-Garb et al., 2023).

The last feature is the conservativity feature. Some deflationist logicians believe that truth is simple because it does not add anything to sentences. In the formal theoretical approach, this means that a theory with the truth predicate should be conservative, as defined earlier. Deflationists have faced objections regarding this point, which will be discussed later in this chapter. Consequently, the next chapter will present some results concerning conservativity in the formal theory of $\mathcal{CT}$.

In summary, deflationism is the position that truth, or the truth predicate, does not contribute to knowledge about the world. The main contemporary positions on deflationism address the question of its commitments. It is important to keep in mind that philosophers tend to focus on these matters in relation to the concept of truth in

---

[3]Following the idea presented in the footnote on page 5, when we say that something characterizes the nature of truth, it means that the corresponding truth theory embodies or represents the concept of truth. The distinction between that and saying that it "completely exhausts the truth predicate" is that it characterizes the concept in that particular way and not in any other way (such as being a generalization device, for example).

natural languages and the world, while logicians maintain this position with respect to the concept of truth in specified languages and formal theories, as discussed earlier. Therefore, it can be understood that there are two approaches to deflationism: one concerning truth in logic and another concerning truth in the world.

4.1   Conservativity in $\mathcal{CT}$

In this section, I will focus on some formal results that are relevant to deflationism. As we saw earlier, one important feature of deflationism is the conservativity of formal theories. That feature has been a matter of debate in relation to some results, especially in the theory of $\mathcal{CT}$. Here, I will present those formal results so that in the next chapters, we can explore different perspectives about the debate and possible solutions to the leading questions in this matter. These results consist, in particular, of $\mathcal{CT}^-$ being syntactically but not semantically conservative over $\mathcal{PA}$ and $\mathcal{CT}$ being non-conservative over $\mathcal{PA}$. They have been proved by (Lachlan, 1981) and appear in the work of (Halbach, 2001; Cieśliński, 2017; Horsten, 2011; Horsten and Leigh, 2017). Here, I will present the main results that concern the investigation about deflationism and provide an overview of each proof's main idea.

For the second theorem, $sub(\phi, \alpha)$ will denote the substitution of numerals for free variables in ϕ in accordance with α, where α is an assignment in M and ϕ is a formula in M. This substitution is performed in M with numerals that might be nonstandard.

Theorem 1 $\mathcal{CT}^-$ *is syntactically conservative over* $\mathcal{PA}$.

This proof is based on a result regarding satisfaction classes; see (Kaye, 1991).

Lemma 1 *All countable models of* $\mathcal{PA}$ *are arithmetically equivalent to a countable model of* $\mathcal{PA}(\mathcal{S})^-$.

Proof idea *(of theorem 1).* Every model of $\mathcal{PA}$ has an arithmetically equivalent model that is expandable to a model of $\mathcal{PA}(\mathcal{S})^-$.

In particular, every model M of $\mathcal{PA}$ has an arithmetically equivalent model M' of $\mathcal{CT}^-$.

Then, if $\mathcal{CT}^- \vdash \phi$, we have that for every countable model $M \models \mathcal{PA}$:

- $M' \models \phi$ and

- if ϕ is arithmetic, then $M \models \phi$.

And then $\mathcal{CT}^-$ is conservative over $\mathcal{PA}$.

Theorem 2 $\mathcal{CT}^-$ *is not semantically conservative over* $\mathcal{PA}$.

This proof is also base on a result in model theory:

Lemma 2 *A countable non-standard model of* $\mathcal{PA}$ *is expandable to a model of* $\mathcal{PA}(\mathcal{S})^-$ *if and only if it is recursively saturated.*

Proof idea *(of theorem 2).* Let M be a countable nonstandard model of $\mathcal{PA}$ such that $(M, T) \models \mathcal{CT}^-$ and let a satisfaction class S be:

$$S = \{(\phi, \alpha) : \phi \in Fm(M) \wedge \alpha \in Asn(M) \wedge sub(\phi, \alpha) \in T\}. \tag{4.1}$$

The set $Asn(M)$ denotes the set of assignments functions in M, so $\alpha \in Asn(M)$ if and only if α is an assignment in M. If we have that $(M, S) \models \mathcal{PA}(\mathcal{S})^-$, then M is expandable to a model of $\mathcal{PA}(\mathcal{S})^-$. So M is recursively saturated. Since there are nonstandard models of $\mathcal{PA}$ that are *not* recursively saturated, thus $\mathcal{CT}^-$ is not conservative over $\mathcal{PA}$.

Theorem 3 $\mathcal{CT}$ *is not conservative over* $\mathcal{PA}$.

Proof idea The essential part of this proof is to show that there is an arithmetic formula that is provable in $\mathcal{CT}$ but not in $\mathcal{PA}$. This will be the formal consistency statement for $\mathcal{PA}$, i.e., $\mathrm{Con}(\mathcal{PA}) = \neg \Pr_{\mathcal{PA}}(\bot)$. We show $\mathcal{CT} \vdash \mathrm{Con}(\mathcal{PA})$ as follows.

First, we need to prove that $\mathcal{CT}$ derives the global reflection principle (GRP):

$$\mathcal{CT} \vdash \forall \psi [Pr_{\mathcal{PA}}(\psi) \rightarrow T(\psi)] \tag{4.2}$$

For showing that we use the fact that $\mathcal{CT}$ proves all the axioms of $\mathcal{PA}$ are true (for this induction is required) and that truth is closed under provability in predicate logic:

33

$$\mathcal{CT} \vdash (Pr_{\mathrm{PL}}(\psi) \rightarrow \psi)^4 \tag{4.3}$$

Given the above, we obtain provability of GRP in $\mathcal{CT}$ by formal induction on the length of proofs of $\mathcal{PA}$.

Finally, the GRP imply:

$$\mathcal{CT} \vdash Pr(\bot) \rightarrow T(\bot) \tag{4.4}$$

from which $\mathcal{CT} \vdash \mathrm{Con}(\mathcal{PA})$ follows using the axioms of $\mathcal{CT}$. And then we obtain that $\mathcal{CT}$ is not conservative over $\mathcal{PA}$.

These results contribute to the ongoing discussions among logicians regarding deflationism. One result, in particular, has sparked objections against deflationism (Shapiro, 1998). However, it is important to note that not all logicians adhere to that critique, with Field being one example (Field, 1999).

Furthermore, we have other results concerning conservativity in various theories. I found the following results with their respective proofs in the work of Cieśliński (2017) and Halbach (2014), who did a very good job compiling and presenting the main results about formal theories of truth. The theories that are not conservative over $\mathcal{PA}$ include $\mathcal{PUTB}$, $\mathcal{FS}$, and $\mathcal{KF}$. On the other hand, $\mathcal{UTB}$ (and consequently $\mathcal{TB}$), $\mathcal{PTB}$, and $\mathcal{FS}^-$ are syntactically conservative over $\mathcal{PA}$.

Regarding semantic conservativity, we have results such as $\mathcal{PUTB}^-$ (and consequently $\mathcal{TB}^-$ and $\mathcal{UTB}^-$), $\mathcal{KF}^-$ with CONS, and $\mathcal{KF}^-$ together with COMPL, which are semantically conservative over $\mathcal{PA}$. However, $\mathcal{TB}$ (and consequently $\mathcal{UTB}$ and $\mathcal{PTB}$), $\mathcal{FS}^-$, and $\mathcal{KF}$ are not semantically conservative over $\mathcal{PA}$.

The reason for focusing on $\mathcal{CT}$ is that, as we will explore in the next chapter, it has been a central point in the literature when discussing the philosophy of truth and deflationism. Nonetheless, examining the conservativity results in other theories will also provide valuable insights.

---

[4] PL denotes classical predicate logic.

Chapter 5

Closing the gap

In this chapter, I will present the perspectives on deflationism from several prominent logicians who also engage with truth from a philosophical standpoint: Field, Cieśliński, Halbach, and Horsten. In the final section, I will offer a critique of these approaches and explore the questions that arise regarding truth, deflationism, and conservativity in both logic and philosophy.

5.1 Revising the literature

It is worth noting that the literature presented in this section is important in the sense that it can be considered as already situated in the middle ground between formal theories of truth and the philosophical aspect. The positions discussed here are concerned with the philosophy of truth in logic and, in that sense, are already narrowing the gap. It is also important to keep in mind that they do not appear to be concerned with truth in the world, which is something I will comment on further in section 5.2. For now, I can only say that this work serves as an entry point with the goal of bridging that existing gap and suggesting a connection between those seemingly distant yet closely related worlds.

5.1.1 Field

In his article "Deflationist Views on Meaning an Content" (Field, 1994), Field explains the contemporary distinction between substantial and insubstantial views on the concept of truth. In "Deflating the Conservative Argument" (Field, 1999), he presents an objection to Shapiro's critique of the conservativity feature and deflationism. According to Shapiro, truth cannot be deflationary if we take into account the results presented in the last chapter (Theorem 3):

> How thin can the notion of arithmetic truth be, if by invoking it we can
> learn more about the natural numbers? (Shapiro, 1998).

35

Field disagrees with this idea. To begin, Field acknowledges that there is no common agreement about deflationism. Nonetheless, he believes that regardless of which version of deflationism one chooses, Shapiro's argument is mistaken. For Field, the main purpose of truth is to enhance the expressive power of language. It serves no other function. Therefore, when Shapiro challenges deflationism with the results concerning conservativity, Field aligns with Parson's response, which posits that what renders a theory non-conservative is not the truth predicate but rather induction. As discussed in the last chapter, $\mathcal{CT}$ is non-conservative, while $\mathcal{CT}^-$ is conservative. Field rephrases Shapiro's question as:

> How thin can the notion of *arithmetic truth* [the italics are mine] be, if by invoking it we can rigorously formulate a more powerful theory about the natural numbers?. (Field, 1999).

Field's response would be that the consequences of truth being a general device are just as uncontroversial as the truth predicate being a generalization device. Furthermore, he would argue that the issue of $\mathcal{CT}$ being non-conservative is related to induction, not truth itself.

> The induction axioms are needed if we are going to derive important facts that involve the notion of truth. (Field, 1999).

Therefore, Field provides a satisfactory response to the concerns among logicians regarding the relationship between deflationism and the results in $\mathcal{CT}$ about conservativity. However, the question then becomes what commitments one must have to hold a deflationist position. As a result, deflationism acquires a more restricted feature in terms of how Field assumes the nature of truth. In other words, Field assumes that the nature of truth in mathematical logic is thin, meaning that it only serves as a generalization device and does not provide any additional function or information to the theories. Based on this, we can conclude that the nature of truth appears to be deflationary within that restricted context, even though there is no conservativity over $\mathcal{PA}$ in the theory of $\mathcal{CT}$.

5.1.2 Cieśliński

In his book, "The epistemic lightness of truth", Cieśliński summarizes contemporary formal and logical results about truth within a philosophical perspective (Cieśliński, 2017). The book is divided into three parts: disquotation, conservativity and reflection principles. In each section, he introduces the formal theories that correspond to each philosophical feature. Therefore, he contributes to bridging the gap between the logical and philosophical approaches.

In the beginning of the first part, Cieśliński explains the meaning of being a disquotationalist. According to him, "Disquotationalists not only claim that *p* and *p is true* are equivalent, but they also claim that equivalences of this sort exhaust the entire content of the notion of truth". He then proceeds to define the disquotational theories, both typed ($\mathcal{UTB}$) and untyped ($\mathcal{PTB}$, $\mathcal{PUTB}$). The last two chapters of the first part focus on the philosophical approaches of disquotationalism. In one of them, he presents formal results about how some of the formal theories can express generalizations, with the conservativity of $\mathcal{TB}$ over $\mathcal{PA}$ being a central result. Cieśliński argues that the reason behind this is that if a deflationist adopts disquotationalism, it is commonly assumed that the extended theory should not bring new results, but he disagrees with this assumption.

> As I take it, the deflationist can respond that proving new arithmetical theorems is a by-product, not a purpose, of increasing expressive power. (Cieśliński, 2017).

Field shares a similar argument to Cieśliński's. For Field, even if the truth predicate can have consequences that imply the non-conservativity of the extended theory, the role of truth is simply that of a generalization device. In the last chapter of the disquotationalism section, Cieśliński discusses a problem with disquotational theories and disquotationalism: some theories are so weak that they cannot even prove the generalizations that a disquotationalist would require. However, Cieśliński provides two solutions for this problem: adding a new rule to the theories (the omega-rule) or adding "psychological dispositions". Although the focus of this section is on conservativity and $\mathcal{CT}$, it highlights another concern of deflationists in logic and how conservativity is still an issue, even if it is not the main focus of this discussion. Therefore, it appears

that conservativity is at the core of the discussions about truth and deflationism.

Moving on to the second part of the book, Cieśliński focuses on conservativity. He starts by presenting the problem with a quote from Horsten:

> The minimal theory entails that a truth predicate should be conservative over a given theory that is stated without the truth predicate (or any other semantic notions). (Horsten 1995, p.183).

Then, he says:

> It is my opinion that any philosophical discussion of conservativity requires a map of the situation in the background; namely, before any deeper philosophical discussion takes place, it should be clear exactly what is at stake in terms of which theories do qualify and which do not. (Cieśliński, 2017).

In Chapters 6, 7, and 9, Cieśliński presents several formal results related to conservativity, including the syntactic conservativity of $\mathcal{PTB}$ over $\mathcal{PA}$ and other theories mentioned in the previous chapter. From Chapter 9 onwards, he shifts his focus to the philosophical aspects of the topic. In Chapter 9, Cieśliński presents his perspective on the conservativity problem from both a semantic and syntactic standpoint. Regarding semantic conservativity, he acknowledges that philosophers like Strollo argue that non-semantic conservativity undermines deflationism. However, Cieśliński disagrees and asserts that there is insufficient evidence to support the claim that deflationists must adhere to non-semantic conservativity as a requirement for their philosophical theory. On one hand, traditional literature on deflationism does not address semantic conservativity, and on the other hand, one should not be concerned about it since the approach is axiomatic.

> Any concept of truth truth that goes beyond its axiomatic characterization is ill conceived at worse and not required at best – that is the story. (Cieśliński, 2017).

In his view, semantic conservativity relies on an understanding of truth that is not strictly deflationary. However, he does acknowledge that it can still be useful in certain philosophical contexts, such as those discussed by Fischer and Horsten. Nevertheless,

he emphasizes that these are more limited projects and do not aim at providing an "overall general theory of truth". (Cieśliński, 2017).

His attitude regarding the syntactic conservativity is different: he thinks that non-conservativity is not an objection against deflationism, but also that it even might help them to defend that approach of the truth predicate. The way to tackle this matter is asking the question about if a proof of a theorem should be either explanatory and (or) justificatory and if the truth predicate play at least of one of those two roles in the formal theories. The reason behind is the criteria that some deflationists, as Horwich, have regarding to truth: truth has to do with explanatory principles, but its main role is to give economy of expressions.

> Substantial scientific concepts do appear in explanations or justifications,
> and they do bring explanatory depth. In contrast, the concept of truth does
> not – and that is the claim. (Cieśliński, 2017).

As we observed in the last subsection, Shapiro criticizes deflationism by arguing that we are "learning more" through the truth predicate from theories such as $\mathcal{CT}$, which are not conservative over $\mathcal{PA}$. For Cieśliński, this implies that the requirement is that the truth predicate is not explanatory or justificatory. The response, then, is that the concepts of explanation and justification are still subjects of debate within mathematics, and it is not clear whether the truth predicate actually explains or justifies something new in formal theories.

> Conservativity means only that another truth-free proof will be available;
> what it does not rule out is that a proof in the extended language will be
> more informative, more general or more explanatory. (Cieśliński, 2017).

Moreover, Cieśliński's position is that syntactic conservativity does not imply the absence of explanatory truth-theoretic proofs, and conversely, non-conservativity does not imply the existence of explanatory truth-theoretic proofs. It is important to note that he is not questioning the validity of the proofs, but rather whether the proofs are justificatory or explanatory.

> To what degree does it justify our belief in consistency of Peano arithmetic
> – that is the question. (Cieśliński, 2017).

Additionally, he seems to agree with Fischer regarding the role of truth, which consists of being a generalizing device. Specifically, the truth predicate helps to shorten proofs. That is why conservative truth axioms assist the deflationist in defending their conception of truth: they demonstrate how its role is fulfilled. What Cieśliński acknowledges here is that this perspective is not entirely aligned with how deflationism was traditionally understood in the philosophical tradition.

In the last two chapters, he explores the question of how useful conservative theories of truth are for the deflationist. However, it is interesting that he revisits the formal results and examines the properties of maximal conservative truth theories. He demonstrates that a maximally conservative theory is syntactically conservative over its base, but it cannot be extended without losing the conservative feature. He believes that the motivation to investigate these theories is that "they describe all the 'light' aspects of truth and these aspects alone" (Cieśliński, 2017). He also presents the positions and responses of both Field and Tennant, but asserts that neither of these two perspectives are convincing.

In the third part of the book, the author offers his point of view in which he thinks he offers a solution both to the first and the second part of the book. He focus on reflection principles and, again, investigates formal results. He shows that:

> An iterated application of reflection principles to the weak typed disquotational theory $\mathcal{TB}^-$ permits us to obtain both the compositional principles and full induction for the language with the truth predicate (Cieśliński, 2017),

and that

> iterated application of reflection principles to the weak untyped disquotational theory $\mathcal{TFB}^-$ permits us to obtain all the axioms of Kripke-Feferman system. (Cieśliński, 2017).

Then, he suggests a response to the question, which he calls 'Q':

> Given that we accept a theory K, why should we accept various additional statements, which are unprovable in K? (Cieśliński, 2017)

And, then, he says that one might think first what are the commitments related to

someone who adheres to a particular theory, he focuses on the concept of belief in a mathematical context and concludes that if the initial theory is "trustworthy", then, we could also trust in the provable statements of a extended theory based on the initial one.

> In particular, the believability theory built over a disquotational theory of truth proves the believability of compositional principles of truth. (Cieśliński, 2017).

This work by Cieśliński is a highly meticulous and valuable presentation of both formal theories of truth and the key features of deflationism within the context of logic. However, it is also important to note the approach he has taken. What I mean by this is that throughout the book, he emphasizes his preference for understanding formal truth theories before delving into the philosophical aspects of deflationism. As a logician, his focus lies in the concept of truth within formal theories rather than natural language, and in fact, he never mentions natural language throughout the book. Additionally, it is interesting to observe that he only briefly mentions scientific theories as a different type of theory. This provides insight into his interests and the guiding principles behind his research.

Nonetheless, as mentioned earlier, Cieśliński explicitly acknowledges the connections between philosophy and logic. It is important to note, however, that he does not provide an explicit definition of deflationism from his perspective. Now, considering his proposed solution regarding conservativity in deflationism, I find that he contributes to the theory with a compelling response by introducing two new and significant concepts to the philosophy of mathematics: explanation and justification. This contribution prompts philosophers and logicians contemplating truth to question how they understand a truth theory in this context. Should a truth theory be explanatory or non-explanatory? Justificatory or non-justificatory? But before delving into these questions, it is crucial to precisely define what we mean by "explain" and "justify". Furthermore, the deflationist must address what they mean by a theory of truth being insubstantial or light. If this implies that a truth theory must not be explanatory and/or not justificatory, then they should be able to respond to whether a truth theory should be conservative or not. And if the answer is affirmative, then the question arises

as to whether formal theories of truth that are not conservative, such as $\mathcal{CT}$, should be disregarded in this context. As demonstrated in this subsection, there are various approaches, but no common agreement has been reached.

5.1.3 Halbach

Halbach, one of the main contemporary logicians working on formal theories of truth, has also expressed his position on deflationism and formal theories of truth. In his book "Axiomatic Theories of Truth" (Halbach, 2014), he introduces axiomatic formal theories and discusses how deflationists can approach each theory. Additionally, in one chapter of the book and in the paper "How Innocent Deflationism is?" (Halbach, 2001), he presents his philosophical stance on deflationism and compositionality.

First, it is important to notice that, for him, the truth axioms involve an ontological commitment, in the opposition of what a deflationist could expect.

> The theory of truth cannot be completely separated from the underlying ontology of objects that can be true, as even very weak axioms for truth will imply that there are at least two different objects, if the axioms imply that something is true and something not true, as even very axioms do. (Halbach, 2001).

One important thing to notice is the introduction of disquotationalism. Here, he explains that the reason why deflationists have rejected the disquotational theories is the deductive weakness. On the other hand, $\mathcal{CT}^-$ and its variants do have that deductive power. But, then, the conservative argument is introduced:

> As McGee (2006) has argued, the deflationist will want to argue that truth is conservative over the base theory without going beyond the base theory, that is, in the present case Peano arithmetic. (Halbach, 2001).

Now, when Halbach focuses on deflationism, he introduces Shapiro's objection and argues against the requirement of conservativeness for deflationism. He distinguishes between conservativeness over logic and conservativeness over the theory of expressions, suggesting that deflationism should not be bound by the former. It is interesting to note that he criticizes Field's distinction between mathematical and truth-theoretical content, while also making a distinction within the inductive axioms as both a syntax

theory and a mathematical theory. Ultimately, this argument leads him to assert that the deflationist should not be concerned about the non-conservativeness of $\mathcal{CT}$.

Finally, I want to add two things: first, that although he does not develop this idea, he suggests that the deflationist should study more closely the typed-free disquotational theories. And, secondly, that he exposes his position regarding truth:

> Deflationism, as I conceive it, can be summed up as follows: truth is only a device of disquotation or a device for expressing 'generalizations' or 'blind ascriptions' - it is neither a property nor a genuine predicate.

Then, based on the definition provided earlier, one can identify Halbach as a disquotationalist. Moreover, despite his focus on the logic and mathematical results of axiomatic truth theories, he makes a significant contribution to the philosophical debate. From his work, we learn that deflationists should approach their position and, specifically, conservativity with caution. Once again, we can observe his concern lies in the nature of truth in formal languages rather than natural languages. However, it is worth noting that he attempts to connect his position with broader philosophical debates, so it remains unclear if his deflationist stance extends to natural language.

5.1.4 Horsten

In "The Tarskian Turn" (Horsten, 2011), Horsten also connects formal theories of truth with deflationism. He raises important questions regarding the compatibility of axiomatic truth theories with deflationism, the commitments of deflationism, and which axiomatic truth theories would be compatible with it. The main idea he presents in this book is what he refers to as "the Tarskian turn". According to Horsten, prior to Tarski, philosophers focused on studying the nature of truth. However, with Tarski's contribution, there was a shift in focus towards understanding how truth functions and is used in specific contexts, such as in axiomatic theories of truth.

According to Horsten (2011), deflationism is a view that considers truth to be a light and insubstantial notion. He also proposes a new version of deflationism called "inferential deflationism". Additionally, he provides his perspective on the classical debate about truth in natural language, stating that substantial theories lack precision. In his view, a scientific theory reflects common sense, and similarly, a formal theory of

truth should extend our common sense understanding of truth. While Horsten devotes a whole chapter to discussing the current state of deflationism in the literature, he also explains the role of disquotationalism and Ramsey's contribution to blind ascription. Furthermore, he acknowledges that a theory of truth does not have to fully capture the meaning of the concept of truth. He also distinguishes between truths and laws of truth, focusing on the latter in his book.

> If it is hopeless to give a substantial definition of truth, then we must try to write down axioms that describe the functioning of the notion of truth as closely as possible—we must try to find the analogues for truth of Newton's laws of force. (Horsten, 1995).

It is worth noting that Horsten is aware of the distinction between natural and formal languages and acknowledges that natural languages are more complex. Interestingly, he justifies the use of formal language by stating that it disregards the complexities of natural languages that are irrelevant for his purposes ((Horsten, 2011)). This observation reveals that when considering the intersection of logic and philosophy in this context, his focus remains on deflationism within formal theories of truth and formal languages, rather than on natural languages and truth in the broader world. The notion of "purposes" can vary, and it would be intriguing to explore those aspects that are currently overlooked.

In some of the chapters, similar to Cieśliński, Horsten introduces the formal apparatus, including preliminaries, disquotational theories, Kripke's theory, and compositional theories. In the latter, he also elucidates his interpretation of the relationship between conservativeness and deflationism. He appears to associate theories that are conservative over $\mathcal{PA}$ with deflationism, while also presenting formal theorems and results. According to him, the insubstantiality of the meaning of truth is the weakness of the truth theory.

Again, it is in the last chapter where Horsten focuses on the connections between the formal theories and the philosophical debate. According to him, an adequate definition of truth can only be provided for a fragment of truth that does not include the truth predicate. Initially, he summarizes the formal theories and argues that the theory $\mathcal{PKF}$ aligns well with deflationism. However, if one chooses to do so, it requires a

re-conception of deflationism itself. The reason behind this is that, in his view, $\mathcal{PKF}$ captures the simplicity of truth, as it can hardly be further reduced. He asserts that it is advantageous that the question revolves around the inferential behavior of truth rather than the nature of truth, as he believes there is no such thing. In his perspective, inferential deflationism conceives the meaning of the concept of truth through rules of inference, of which axioms are limited cases. Furthermore, truth is considered insubstantial in the sense that it "cannot be described in terms of unrestricted general laws", but only through restricted laws of truth. This distinction sets truth apart from other mathematical and scientific concepts, such as properties and relations.

> So, one can take the fact that truth theories allow us to simulate fragments of analysis to show that truth theories introduce substantial new ontological commitments and are thus inflationary. But the message can also be taken to be that this fact shows that fragments of analysis can be interpreted as carrying no extra commitments over arithmetic and can thus be interpreted in a deflationary way. The latter is in the spirit of Feuerman's way of looking at the matter.

> Let us finally return to the lives and times of Galilei and Newton. Galilei transformed our conception of force. The mathematization of the concept of force was a way of deflating the concept of force. But heated discussions about the nature of force persisted long after his death and continue on to this day. I suspect that for truth, the situation is no different. Tarski's logical analysis of the concept of truth is deflationary in a way that we found hard to express accurately and precisely. (Horsten, 2011).

It is interesting and significant to observe that Horsten interprets the Tarskian turn in these terms, drawing a parallel to the turn in the history of physics with Galileo and Newton and their concepts, such as the concept of force, as illustrated in the aforementioned quote. This demonstrates his influence from the philosophy of science and physics and how it informs his approach to the philosophy of truth in this particular context.

> The Tarskian turn will not make the discussion about the nature of truth go away. Perhaps this is as it should be. Foundational discussions about

the nature of force have in the long run been highly beneficial for the development of theoretical physics: think about how the discussion about actio in distans stimulated the development of the field concept. Let us hope that the continuing discussion about the nature of truth will turn out to be as fruitful in the long run. (Horsten, 2011).

5.2   Is truth deflationist?

In the chapter "What is deflationism?" (see page 29), I explained the philosophical position of deflationism and how it was understood in the literature in relation to truth in the world and truth in logic and mathematics. Although there is no common agreement on the exact meaning of deflationism, we can create two labels to make the concept more concrete: $Deflationism_W$ and $Deflationism_L$. The former refers to the position that the concept of truth is not substantial when used in natural languages and can include metaphysical commitments regarding the existence of (pro)sentences and their relationship to truth. The latter refers to the position that truth is not substantial within mathematics and mathematical logic, particularly when studying formal theories of truth.

My aim in the following is to bridge the gap between these two perspectives using these new terms. This will involve three main points: first, whether axiomatic formal theories of truth can be used by philosophers who defend $Deflationism_W$; second, whether logicians and philosophers who defend $Deflationism_L$ could use a broader context in the philosophy of truth; and third, whether logicians who focus solely on formal theories of truth could explore other positions to inspire their theories.

In the first place, it is important to mention that what I am about to discuss is related to the distinction between the truth predicate and the nature of truth. As some philosophers (like Jago) and logicians have clarified, the focus of the literature tends to be on the truth predicate rather than the concept of truth in general or the nature of truth. The question about the nature of truth refers to the question "What is truth?" and is therefore related to the concept of truth. All of this is to say that there is a relationship between the question about the truth predicate in a restricted and formal language in logic and a more general question about the truth predicate in natural

language and/or the question about the concept of truth or the nature of truth in natural language. Thus, when someone that adheres to $Deflationism_L$ talks about truth in that restricted language, they are also talking about the nature of truth and the concept of truth in general. However, it is very important to keep these distinctions in mind.

In this sense, philosophers who adopt the position of $Deflationism_W$ should also consider formal results as part of the nature of the world in general. This leads to further questions about how to connect formal languages with natural languages. As Horsten mentioned, when working solely in mathematical logic, we get rid of ambiguity and other features that are not useful for mathematical objectives. However, when one focuses on the nature of truth, the question about the relationship between those two approaches of languages becomes important, and the ambiguity that was eliminated in the formal language could possibly lead us to some knowledge in our investigations about the nature of truth.

Therefore, the one who adheres to $Deflationism_W$ can argue, if we accept the arguments of Field, Halbach, Cieśliński, and Horsten, that at least in that part of the world, the truth is deflationist. However, this does not imply that it is true for the world in general.

On the other hand, the logicians and philosophers who adheres to $Deflationism_L$ seem to ignore or omit the classical and contemporary literature on the nature of truth. They claim to be interested in the concept of truth, but it is not clear why they choose to focus only on the truth predicate in those specific languages and why they do not investigate the nature of truth in natural language. Although it seems understandable why one would start there (it is easier to work with formal languages, as Horsten noted, since they are more precise and omit ambiguity), it is still essential to make it explicit so that both approaches (the narrow concern about truth in formal theories and the broader concern about truth in natural languages) can collaborate with each other. However, it is still interesting to remember that Tarski did make explicit the distinction and why he was not focusing on truth in natural languages: he was pessimistic about that investigation since he had a positivist belief in which he only believed in the scientific method and thought that the matter of truth in natural language could not be treated with enough rigor.

We now turn to the question of whether the concept of "truth" in natural languages can aid the formal theories. The last paragraph is particularly interesting in this regard, because some logicians (like Horsten) argue that the formal theories reflect our common sense or understanding of truth. However, it seems to me that the theories are motivated by certain intuitions rather than what people generally conceive as truth. For example, Kripke's semantic theory reveals a very different intuition about truth, so it would be intriguing to explore how other positions could motivate alternative formal theories. Nonetheless, it is difficult to see beyond how these theories can be applied, since current investigations in the truth theories are concerned with conservativity in other branches of mathematics. And, in any case, it is important to notice that the $Deflationism_W$ are influencing in some sense the formal theories in the sense that they are taking that deflationist intuition to build the theories, one can confirm that in the axioms and rules of inference. So, it becomes an interesting relation in which if we start by those intuitions, then we work on the formal theories (like $\mathcal{CT}$) and then the $Deflationism_L$ try to defend that which formal theories are a good characterisation for the deflationist position.

Taking into account the last sentence, it is important to note that theories like $\mathcal{CT}$ do accurately characterize $Deflationism_L$. As previously shown, the philosophers and logicians defending this connection respond well to objections. However, it is important to remember that for these responses to be more satisfactory, the investigation of the concepts of justification and explanation in mathematics needs to be deepened. This would enable us to draw more conclusive results.

In light of these ideas, I would like to add, in accordance with the first footnote in this text, an assessment or suggestion regarding a question that frequently arises in the literature. If we adhere to Horsten's suggestion, we may become aware of the "Tarskian turn". This involves refraining from asking which formal theories better characterize the deflationist positions and instead focusing on studying how truth behaves within the formal theories. We can examine whether truth exhibits deflationist characteristics within these theories and also revisit the other classical theories mentioned in the chapter "Truth in the World" to investigate if they behave differently (such as in a coherent manner, for example).

From what we can conclude so far, truth appears to be like $Deflationism_L$ claim, but the jump to $Deflationism_W$ is not yet feasible. I would like to emphasize that even if truth appears to be like in $Deflationism_L$, it would still be productive for $Deflationism_L$ to consider different levels at which truth is non-substantial. Depending on the chosen truth theory, and whether or not it is conservative, we can see that the nature of truth functions in different ways. From the philosophers who worked on truth in natural languages, we may learn and consider a position that views the nature of truth as less heterogeneous, the pluralists in the Modern Debate (see in 10). This is supported by a Horstian position, in which the main question is how truth behaves in the world, and specifically, how it behaves in formal theories. In other words, each formal theory contributes to our understanding and knowledge of the nature of truth by demonstrating its behavior. As we can see, in most cases, it is deflationist, and in the cases where the theories are not conservative, the investigation is still ongoing, but it is still motivating for both $Deflationism_L$ and $Deflationism_W$ approaches regarding further research and results concerning the concepts of explanation and justification, and the role of the truth predicate in proofs.

In summary, the response to the question of whether $\mathcal{CT}$ characterizes deflationism appears to be positive. However, there are still unanswered questions regarding the concepts of justification and explanation. Additionally, it would be interesting to explore a pluralist position, particularly in relation to the debate on conservativity. Now, regarding the three points I mentioned earlier:

Firstly, I suggest that the investigations conducted in $Deflationism_L$ should be considered as a step forward in $Deflationism_W$ since formal theories are a part of the world. Secondly, even though the question of whether $\mathcal{CT}$ characterizes deflationism seems to be well-founded, I propose that philosophers and logicians who are concerned with the nature of truth in formal theories also consider the other classical positions, such as pragmatism and coherence. Following Horsten's idea, the question should focus on how truth behaves in formal theories and how it can be related to deflationism, as well as other theories. Thirdly, I believe it is possible for logicians who are solely focused on formal theories to explore other philosophical positions in order to inspire their theories.

Chapter 6

Conclusions

The aim of this work was to understand the concept of truth through the deflationist position and its relation to axiomatic theories of truth, specifically compositional truth and the feature of conservativity; and to bridge the gap between the philosophers that worry about truth in the world (recall chapter 2) and the logicians that worry about truth in logic (recall chapter 3). I explored the relationship between the logical, formal, and philosophical perspectives. As mentioned in the beginning, the distinction can become blurred, but it remains useful when we consider the differences in methodologies and languages used to study the truth predicate.

When the question about the nature of truth was discussed in the second chapter, which can also be considered as part of the more philosophical literature, it was found that in the classical debate there are, in principle, five positions regarding the metaphysical status of truth. In the modern debate, on the other hand, there are two positions, both of which have a version of deflationism. In the classical debate, the positions were: the correspondent theory of truth, coherence theory, pragmatist theory, redundant theory, and Tarskian theory of truth. The correspondent theory tries to connect sentences with objects or facts in the world through an isomorphism. The coherence theory states that the nature of truth is mainly connected to how coherent it is with other beliefs and/or propositions. The pragmatist theory places more emphasis on the usefulness of truth, suggesting that the nature of truth depends on its utility conditions. The redundant theory (which corresponds to deflationism) is commonly accepted in the literature, starting with Ramsey, and asserts that the truth predicate is redundant and does not add much information to sentences. Finally, the Tarskian theory adopts the Tarski biconditionals as an explanation of the nature of truth. On the other hand, the modern approach is based on truth-bearers and focuses on the question of whether truth is substantial or not. Truth-bearers can be facts, moments, or similar concepts, and deflationism corresponds to the view that the nature of truth is insubstantial.

In contrast, logicians have primarily focused on formal languages when contemplating the concept of truth. The distinction between an object-language and a meta-language has been central in their considerations. Additionally, we have seen that there are two ways to approach truth from a logical perspective. The first one is called semantical theories of truth. We briefly explored how Tarski and Kripke defined a notion of truth for specific models.

On the other hand, the axiomatic approach aims to characterize the truth predicate through a set of axioms rather than providing a definition. The axiomatic theory of $\mathcal{CT}$ was defined. At this point, the notion of conservativity was introduced with the purpose of relating it to the question of how deflationism could be connected to these theories. If we adopt a deflationist approach to truth, then the selected theory should be conservative over the object-theory, or not.

The motivation for this distinction regarding how truth has been studied was explained. After that, it was explained that although the deflationism position is not precisely defined, there are some main points in the contemporary literature. These include the idea that the nature of truth is thin, that the Tarski-biconditionals exhaust all the meaning of truth, that the utility of truth consists of making generalizations, and that truth theories should be conservative.

Then, some results about conservativity in $\mathcal{CT}$ were presented. Finally, in the chapter titled "Closing the gap" (see page 35), there was a close examination of the logicians and philosophers who advocate for a $deflationist_L$ position. They are actively closing the gap by addressing the philosophical question about the nature of truth within the framework of formal theories. It is evident that they are concerned about the non-conservativity of certain theories, such as $\mathcal{CT}$, and they provide responses to those objections. The initial gap can be divided into two parts: the philosophical and logical approaches to the question of the nature of truth, and the philosophical approach to the nature of truth in natural languages. The distinction between two types of deflationism, as mentioned earlier, is seen as an initial step towards closing the gap. $Deflationism_L$ occupies a middle ground in this regard. However, the transition from $Deflationism_L$ to $Deflationism_W$ is still unresolved and remains an open question.

In the last section, "Is truth deflationist?" (see page 46), I explained how I be-

lieve that $Deflationism_L$ has done a commendable job in defending its position against objections, particularly concerning conservativity. This defense has given rise to new unanswered questions pertaining to the concepts of explanation and justification. However, I also underscored the importance of distinguishing between $Deflationism_L$ and $Deflationism_W$, which had not been previously made and is pivotal to this discussion. In conclusion, we can acknowledge that although $Deflationism_L$ is well defended against objections, it does not encompass a more comprehensive $Deflationism_W$. This exploration into the nature of truth in formal theories and its philosophical underpinnings contributes to the broader investigation of truth in the world and in natural languages, as these theories are regarded as part of the world and provide partial descriptions of how truth operates within it. Consequently, further inquiry into the relationship between natural languages and formal and mathematical languages is warranted.

In summary, the response to the question of whether $\mathcal{CT}$ characterizes deflationism appears to be positive. However, there are still unresolved issues concerning the concepts of justification and explanation. Moreover, considering a pluralist position, particularly in relation to the conservativity debate, would be an intriguing avenue to explore. Now, let's revisit the three points I mentioned earlier, with an additional one:

Firstly, I propose that the investigations carried out in $Deflationism_L$ should be viewed as a progression towards $Deflationism_W$ since formal theories are an integral part of the world. Secondly, while the question of whether $\mathcal{CT}$ characterizes deflationism seems well-grounded, I suggest that philosophers and logicians who delve into the nature of truth in formal theories also take into account other classical positions, such as pragmatism and coherence. Following Horsten's notion, the inquiry should focus on how truth behaves in formal theories and how it can be connected to deflationism, as well as other theories. Thirdly, I believe that logicians who primarily concentrate on formal theories have the potential to explore other philosophical positions as a source of inspiration for their theories. And lastly, the leap from the philosophy of truth in logic to the philosophy of truth in the world remains open, but it is both highly interesting and worth considering.

Bibliography

Aristotle (1924). *Aristotle's Metaphysics*. Clarendon, Oxford, 1. ed. edition.

Armour-Garb, B., Stoljar, D., and Woodbridge, J. (2023). Deflationism About Truth. In Zalta, E. N. and Nodelman, U., editors, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Spring 2023 edition.

Betti, A. (2015). *Against facts*. MIT Press.

Cieśliński, C. (2017). *The epistemic lightness of truth: Deflationism and its logic*. Cambridge University Press.

de Nos, C. (2012). Lo malo de ser bueno. Álbum "Porfiado". Lyrics and music by Roberto Musso.

Dummett, M. (1976). What is a theory of meaning?(ii).

Field, H. (1994). Deflationist views of meaning and content. *Mind*, 103(411):249–285.

Field, H. (1999). Deflating the conservativeness argument. *The Journal of Philosophy*, 96(10):533–540.

Fischer, M., Halbach, V., Kriener, J., and Stern, J. (2015). Axiomatizing semantic theories of truth? *The Review of Symbolic Logic*, 8(2):257–278.

Frápolli, M. J. (2005). Ramsey's theory of truth and the origin of the pro-sentential account. *FP Ramsey. Critical reassessments. Continuum*, pages 113–138.

Frápolli, M. J. and Nicolás Marín, J.-A. (1997). Teorías de la verdad en el siglo xx.

Frege, G. (1918). Thoughts.

Frege, G. (1956). The thought: A logical inquiry. *Mind*, 65(259):289–311.

Glanzberg, M. (2021). Truth. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Summer 2021 edition.

Halbach, V. (2001). How innocent is deflationism? *Synthese*, 126(1/2):167–194.

Halbach, V. (2014). *Axiomatic theories of truth*. Cambridge University Press.

Halbach, V. and Leigh, G. E. (2023). *The Road to Paradox: a guide to syntax, truth*

and modality. To be published, 1. ed. edition.

Hodges, W., Wilfrid, H., et al. (1993). *Model theory*. Cambridge university press.

Horsten, L. (1995). The semantic paradoxes, the neutrality of truth and the neutrality of the minimalist theory of truth.

Horsten, L. (2011). *The Tarskian turn: Deflationism and axiomatic truth*. MIT press.

Horsten, L. and Leigh, G. E. (2017). Truth is simple. *Mind*, 126(501):195–232.

Jago, M. (2018). *What truth is*. Oxford University Press.

Kaye, R. (1991). Models of peano arithmetic.

Lachlan, A. H. (1981). Full satisfaction classes and recursive saturation. *Canadian Mathematical Bulletin*, 24(3):295–297.

Marker, D. (2006). *Model theory: an introduction*, volume 217. Springer Science & Business Media.

Mosquera Restrepo, J. (2022). Si, la verdad las paredes hablaran. comisiondelaverdad.co.

Mulligan, K., Simons, P., and Smith, B. (1984). Truth-makers. *Philosophy and phenomenological research*, 44(3):287–321.

Picollo, L. (2020a). Alethic reference. *Journal of Philosophical Logic*, 49(3):417–438.

Picollo, L. (2020b). Reference and truth. *Journal of Philosophical Logic*, 49(3):439–474.

Ramsey, F. P. (1991). *On truth : original manuscript materials (1927-1929) from the Ramsey Collection at the University of Pittsburgh*. Episteme. Kluwer Academic Publishers, Dordrecht ;.

Ramsey, F. P. and Moore, G. E. (1927). Symposium: Facts and propositions. *Proceedings of the Aristotelian Society, Supplementary Volumes*, 7:153–206.

Shapiro, S. (1998). Proof and truth: Through thick and thin. *The Journal of Philosophy*, 95(10):493–521.

Szabó, Z. G. (2022). Compositionality. In Zalta, E. N. and Nodelman, U., editors, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford Univer-

sity, Fall 2022 edition.

Tarski, A. (1944). The semantic conception of truth: and the foundations of semantics. *Philosophy and phenomenological research*, 4(3):341–376.

Tennant, N. (2005). Deflationism and the gödel phenomena: Reply to ketland. *Mind*, pages 89–96.

Walker, R. C. (2017). Theories of truth. *A Companion to the Philosophy of Language*, pages 532–555.

Wittgenstein, L. (2009). Tractatus logico-philosophicus.–wittgenstein l. *Izbrannye raboty*, pages 10–227.